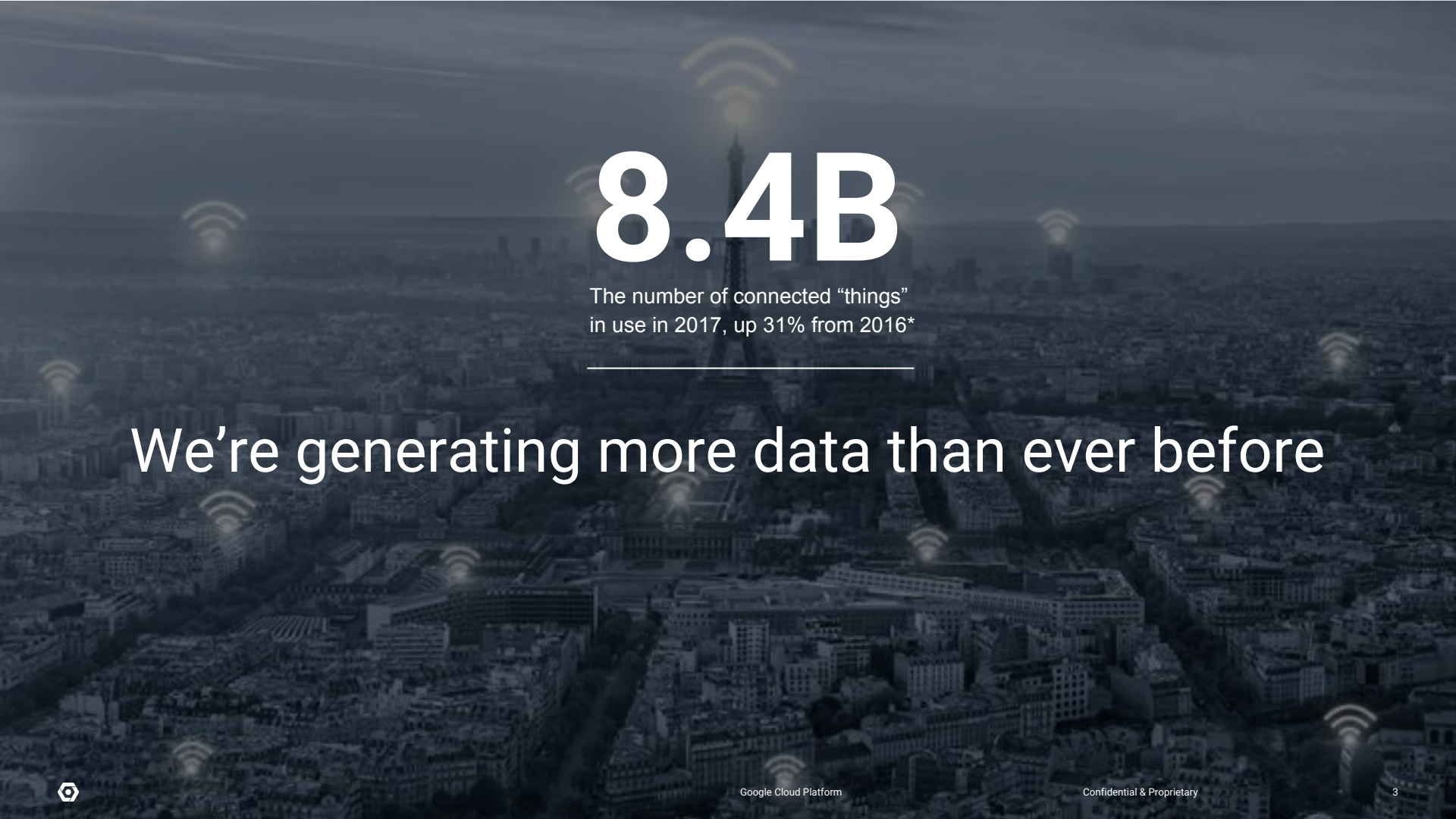
IoT at Google Cloud

Preston Holmes
Head of IoT Solutions

Google Cloud Platform

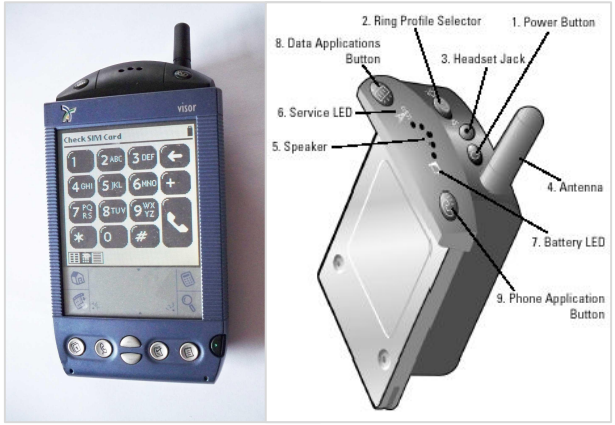
Organize the world’s information and make it universally accessible and useful.

Google’s Mission


8.4B
The number of connected “things”
in use in 2017, up 31% from 2016*
We’re generating more data than ever before

IoT is a period of transformation

Phone

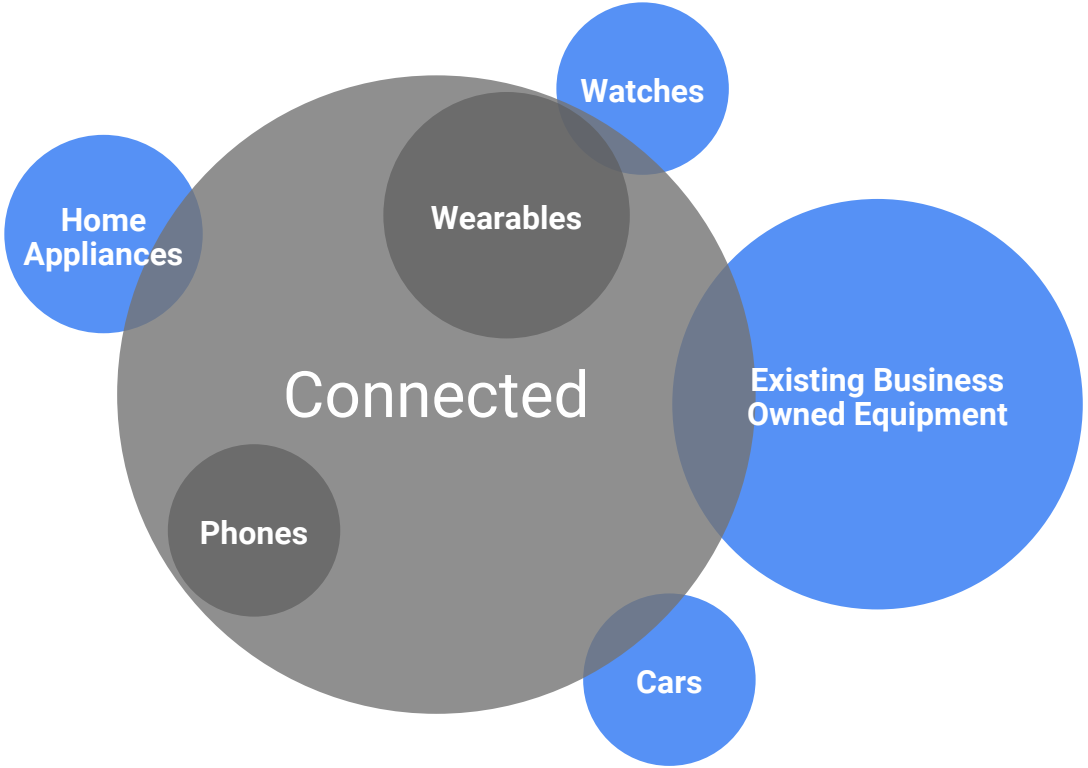


"IoT"

Phone

IoT is a period of transformation

Not Connected



Information is Everywhere But it’s not DATA Yet

Home

Cities

Retail

Transportation

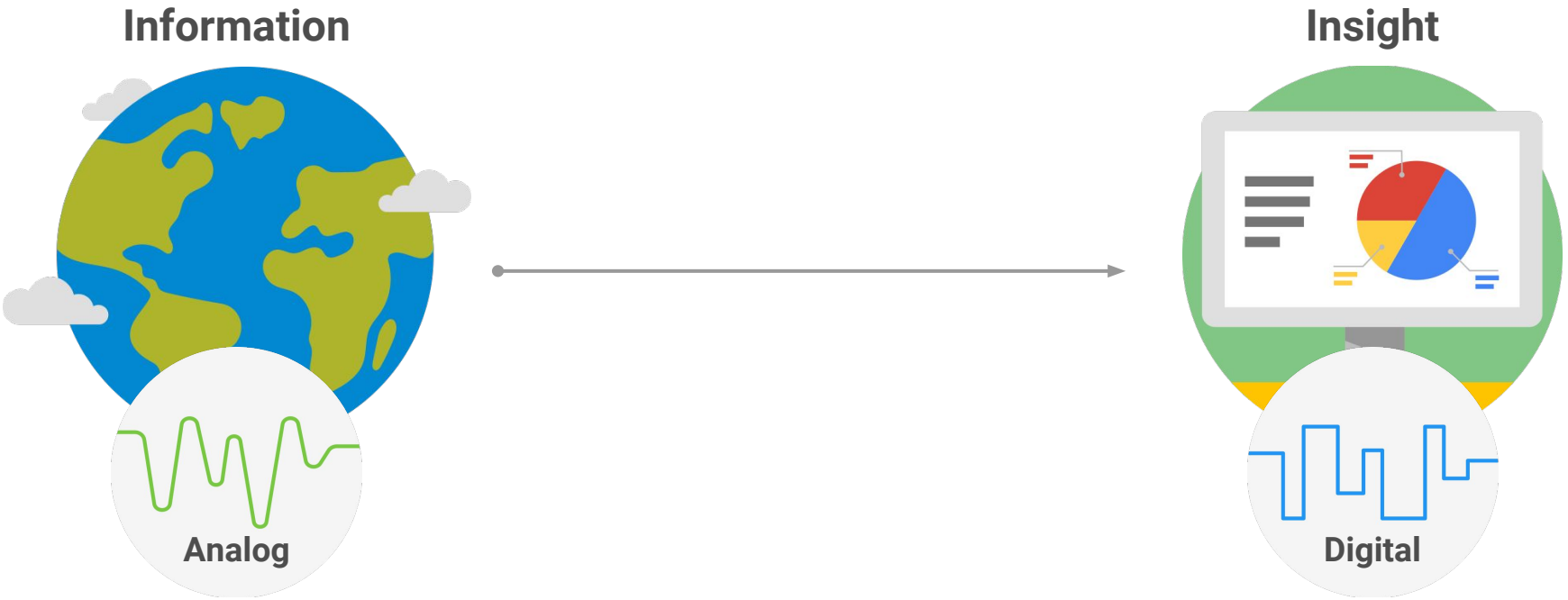
Buildings

Manufacturing & Industrial

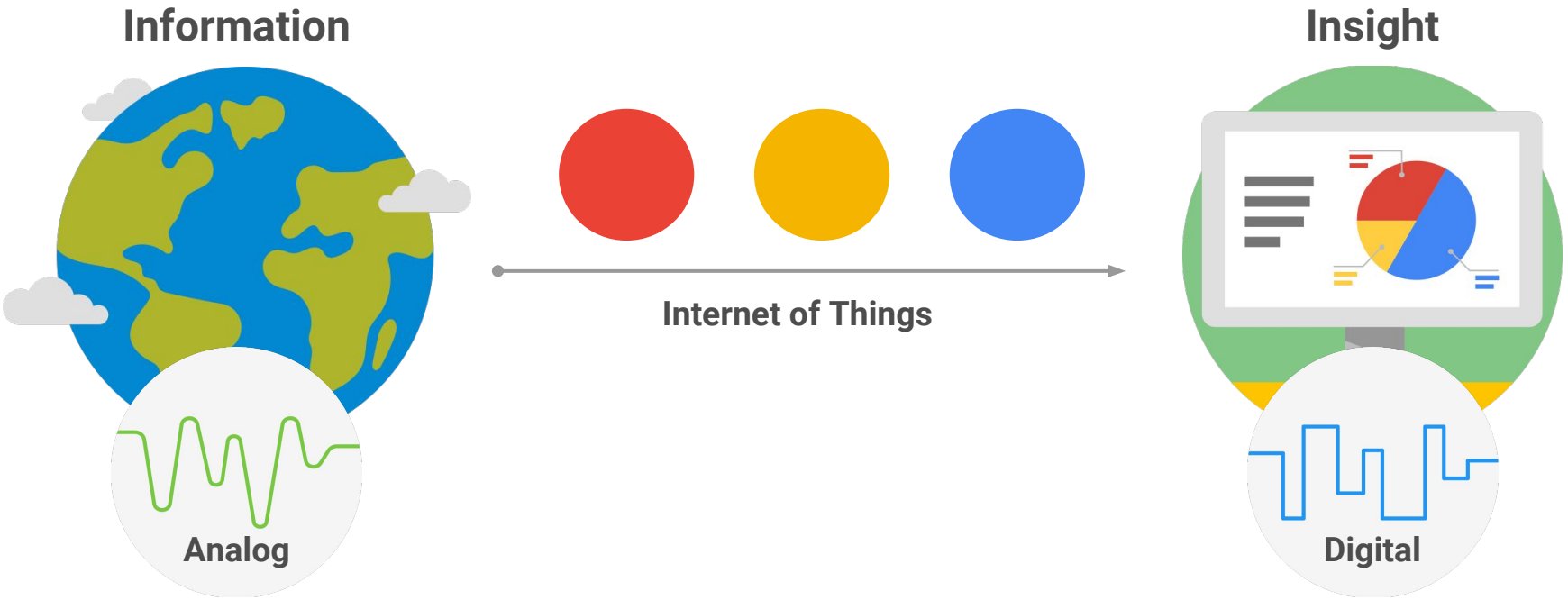
Healthcare

How do you collect and process this analog information, to transform into useful business Intelligence?

How do you collect and process this analog information, to transform into useful business Intelligence?



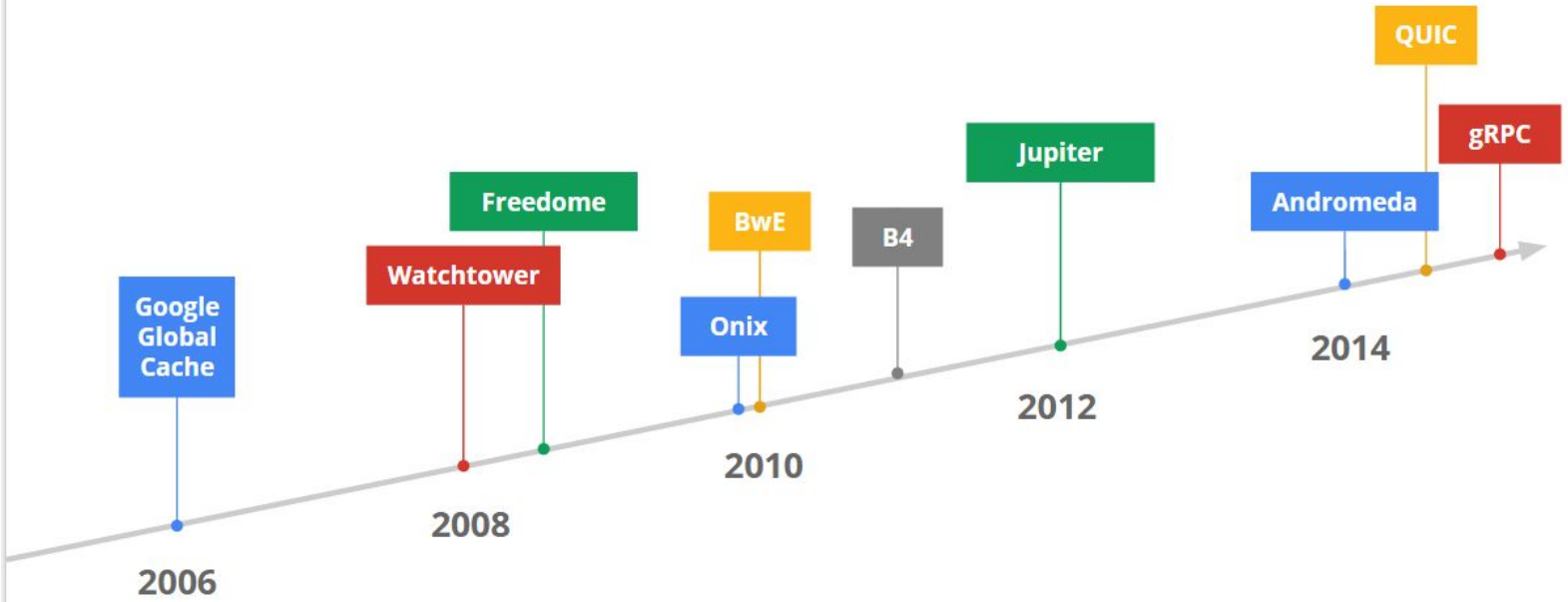
How do you collect and process this analog information, to transform into useful business Intelligence?




A datacenter is not a collection of computers,
a datacenter is a computer.

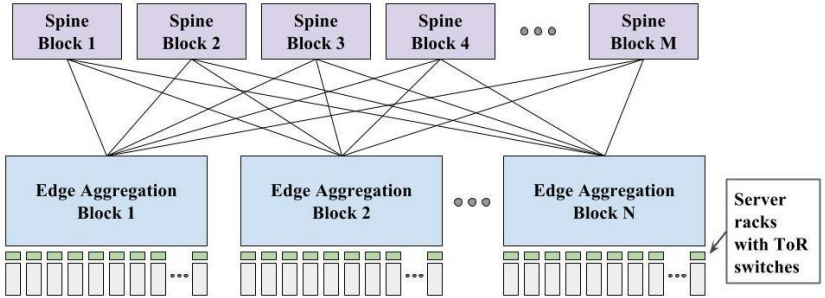

Laying undersea cable

Google Innovations in Networking



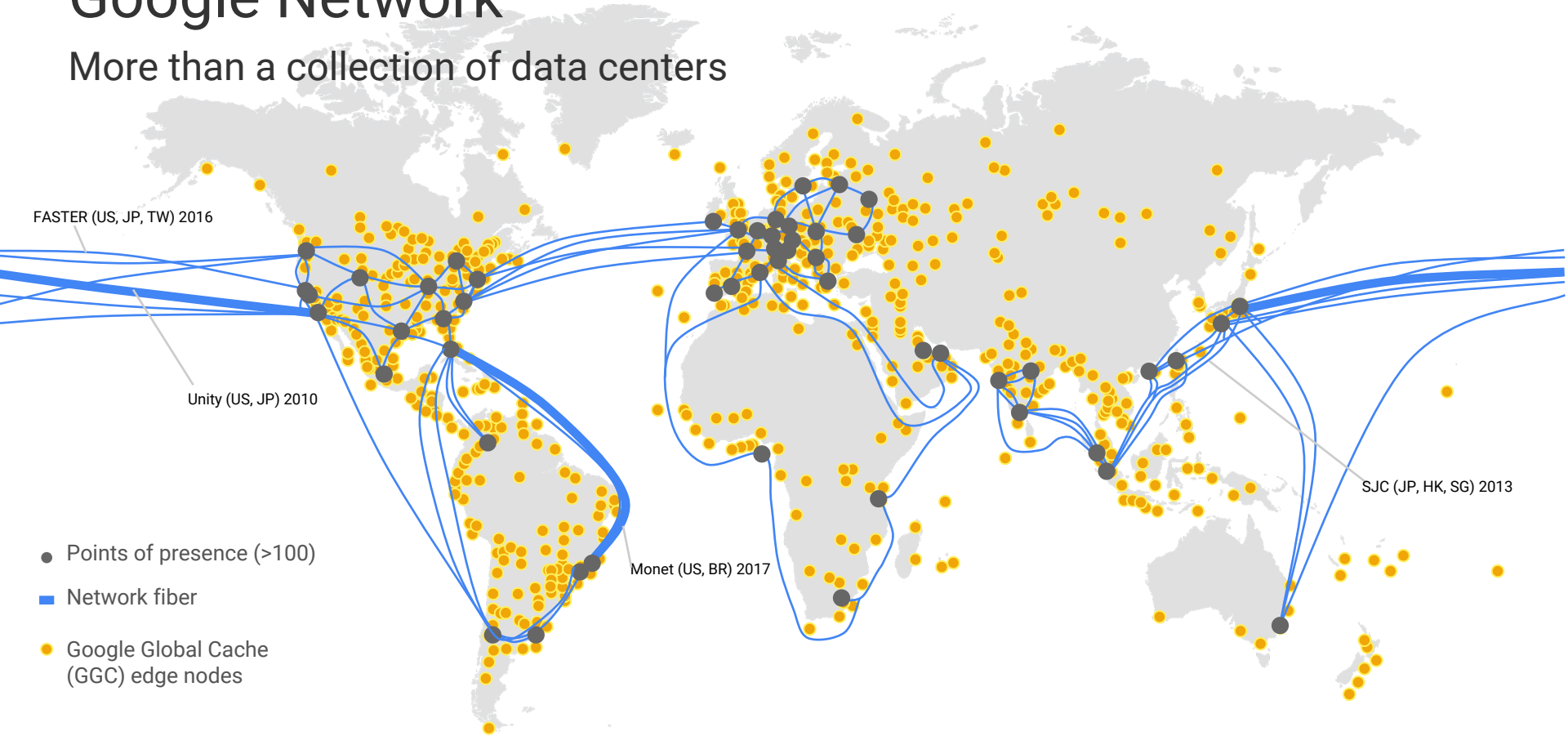
Google Jupiter:

On the left is a Jupiter superblock. It is one part of a fabric which interconnect up to 100,000 servers at 10 Gbit/s each—more than 1 Petabit/sec of total bisection bandwidth for transfer of information between physical and virtual machines It's enough to transfer the entire scanned contents of the library of congress in 1/10 of a second.

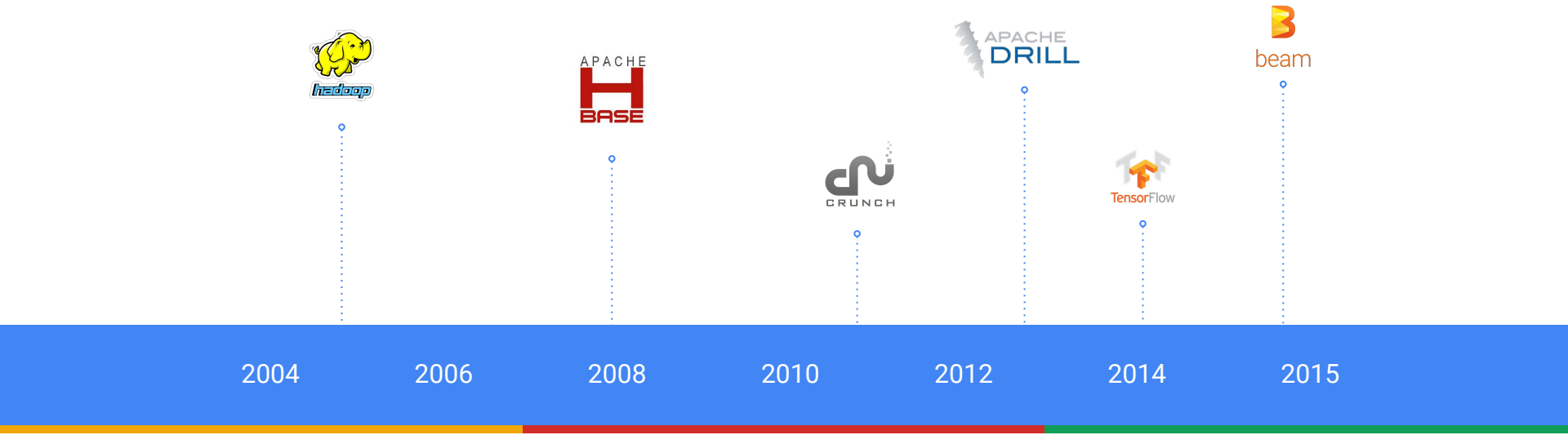


Google Network

More than a collection of data centers

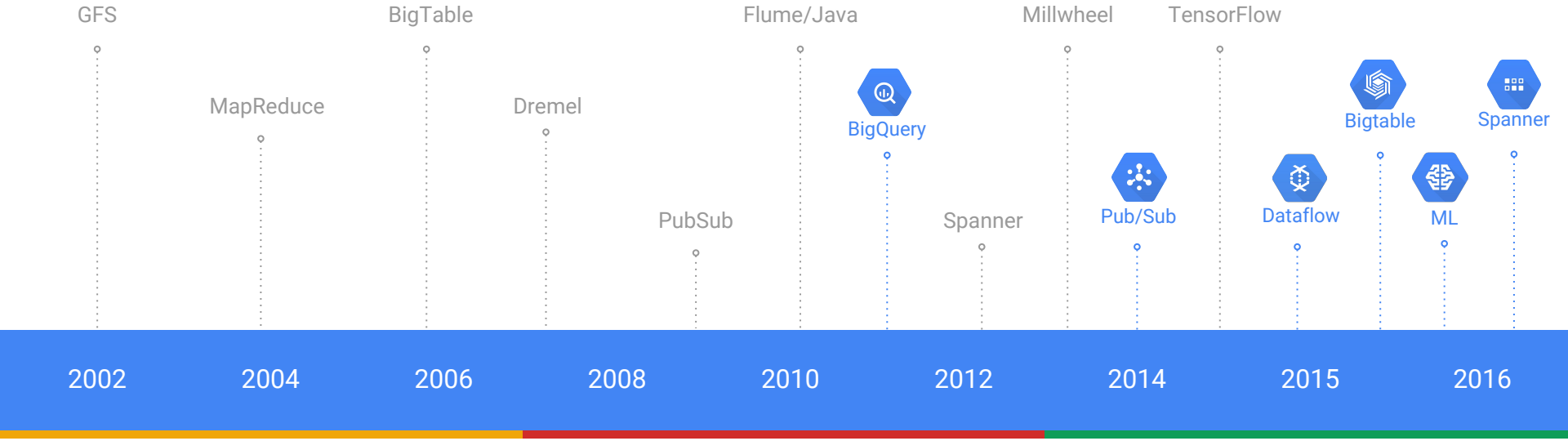


Our research and legacy in data management and analytics stack run deep



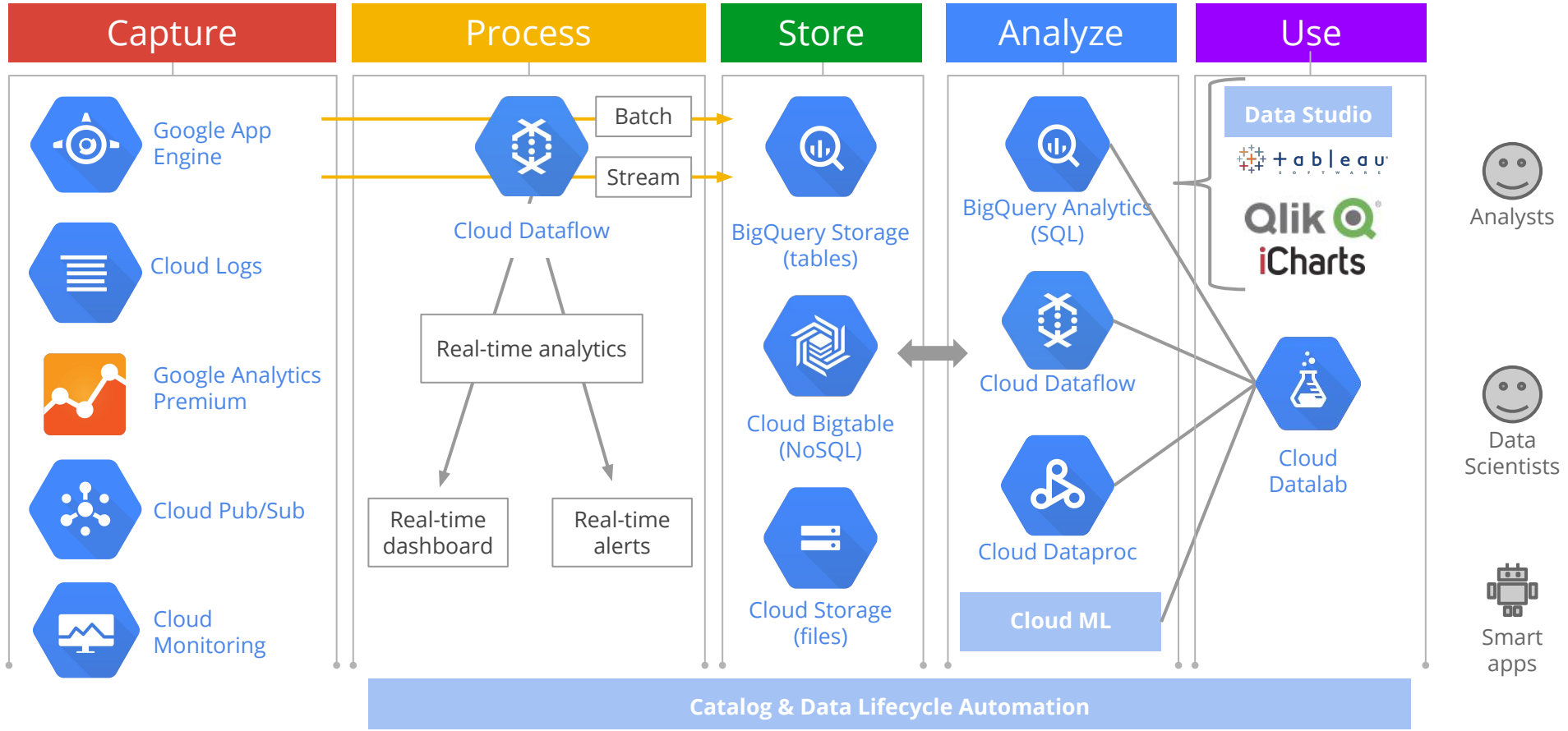
Open Source

Our research and legacy in data management and analytics stack run deep

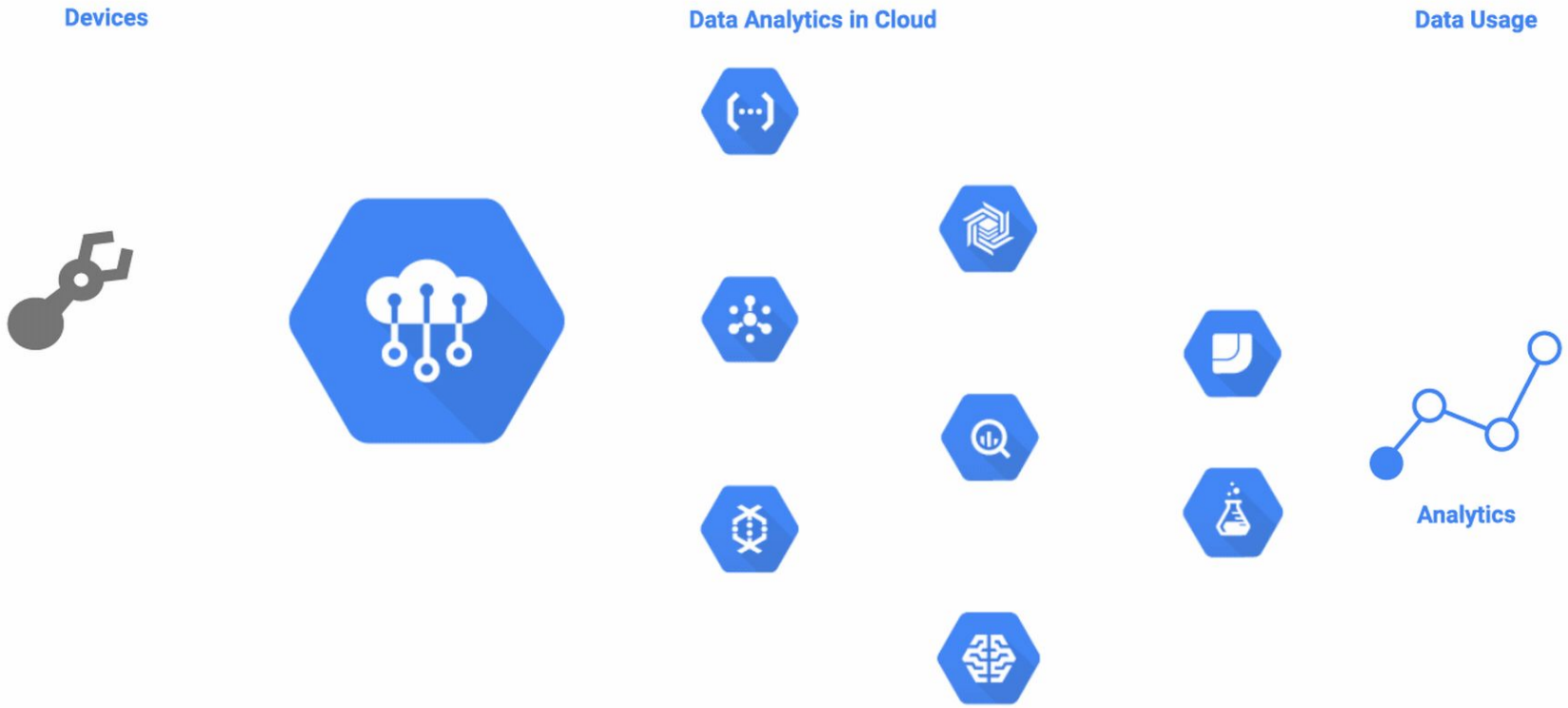


Google Papers + Products

Google Big Data Pattern



IoT Core handles device management and bi-directional device communication as part of an over all GCP IoT Solution



What Services are in IoT-Core ?

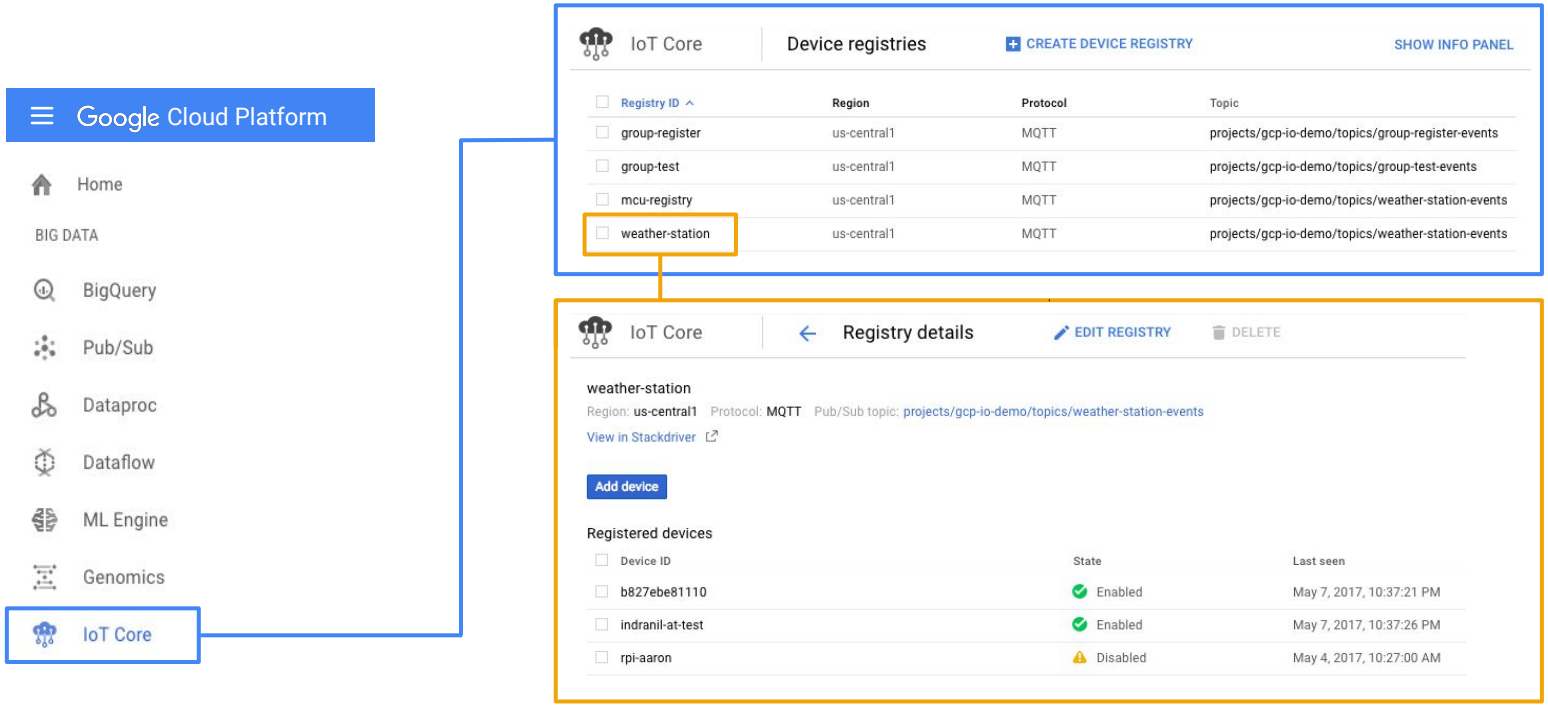
Device Manager

- Maintains registries of devices as named logical resources
- Protects these entries with IAM permissions
  - eg who can delete a device
- Associates device credentials with these resources
- Acts as the identity provider (IdP) for the MQTT broker
- Provides storage and delivery API for device configurations
- Maintains some operational state metadata for the device:
  - enabled/disabled
  - connectivity and error status

MQTT Broker

- Provides a stateful socket connection to devices for bidirectional communication
- Brokers device->cloud telemetry data onto Cloud PubSub for use in downstream GCP products or customer applications
- Delivers configuration updates via a Device Manager API
- Is exposed through a global DNS endpoint over multiple ports

Simple UI for monitoring and management



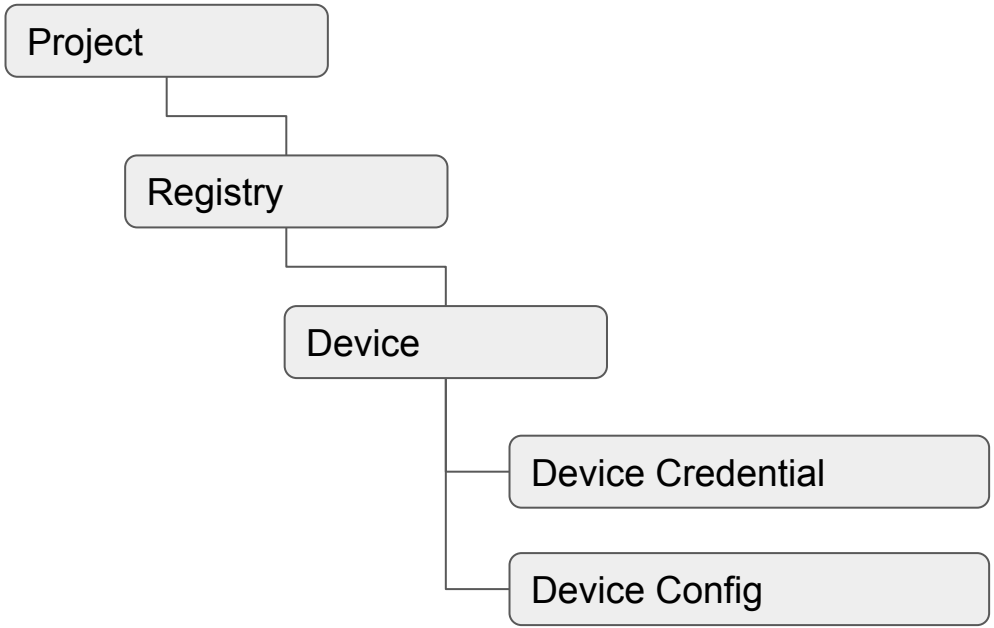
What is MQTT

MQTT is a machine-to-machine (M2M)/"Internet of Things" connectivity protocol. It was designed as an extremely lightweight publish/subscribe messaging transport.

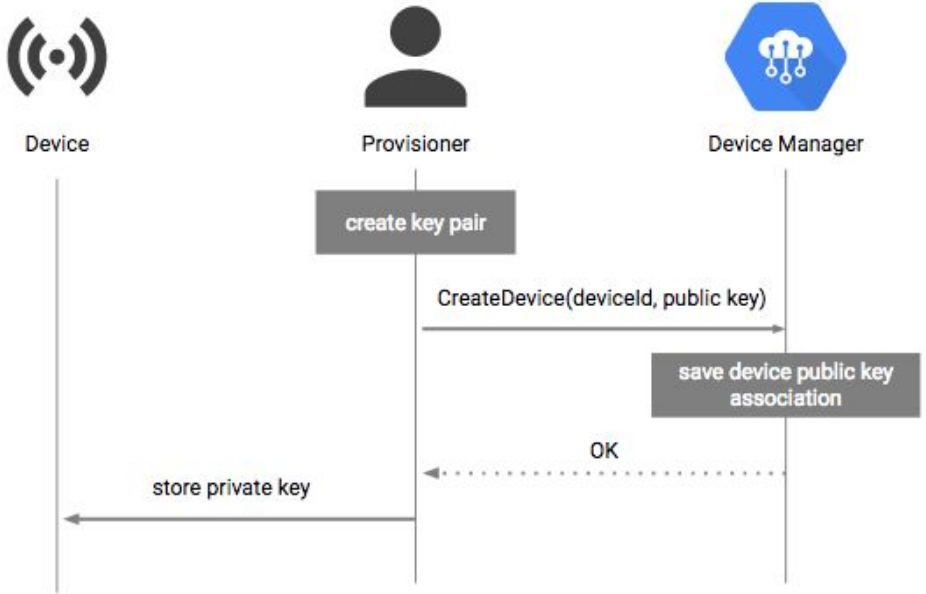
- Originally developed at IBM in 1999 and designed for constrained devices
- Is now a ratified standard
- Is a binary protocol making efficient use of over-the-wire bandwidth
- Is simple to implement and so has many small memory footprint libraries available
- Has become a common and de-facto standard used in many IoT projects

Device Manager Resource Model

The device manager organizes cloud resources to handle device management.



Device Manager: Devices : Identity



Device identity is based on an asymmetric key-pair of two supported formats:

- RSA 256 public key wrapped in a X.509v3 certificate
- Elliptic curve (ECDSA) algorithm using P-256 and SHA-256 [more efficient, better suited for small devices]

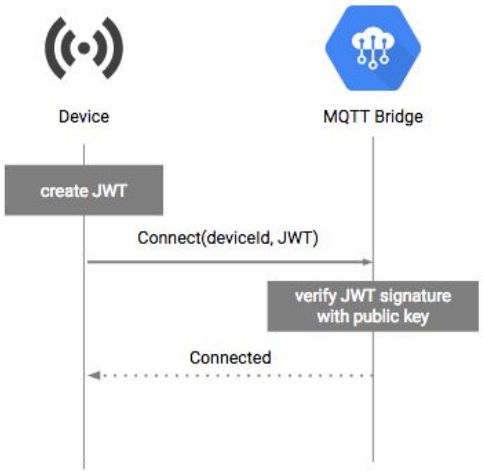
Credentials may optionally have an expiration timestamp

A device can have up to 3 credentials associated with it at a time, allowing for rotation

The service should never need the private key

The sequence shown here is only one way to handle device provisioning

MQTT Broker: authentication



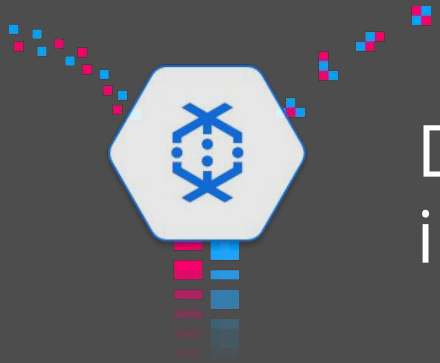
A device authenticates with two pieces of information:

1. the MQTT client ID (a feature of the MQTT protocol) which must be in the form of device name:

   projects/{project-id}/locations/{cloud-region}/registries/{registry-id}/devices/{device-id}

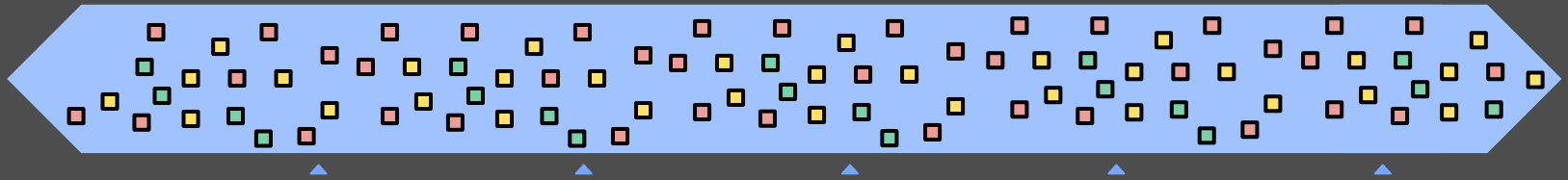
2. An MQTT password in the form of a JWT token signed by the device's private key

- The "username" field in MQTT clients is ignored
- JWT token may have a max expiration of 1 hour
- Device's clock must be within 10 minutes of Google's time (use Google NTP)

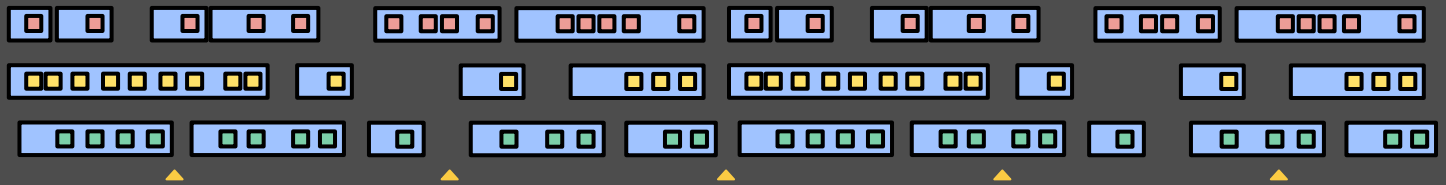


Dataflow: organize torrents of IoT data into actionable windows

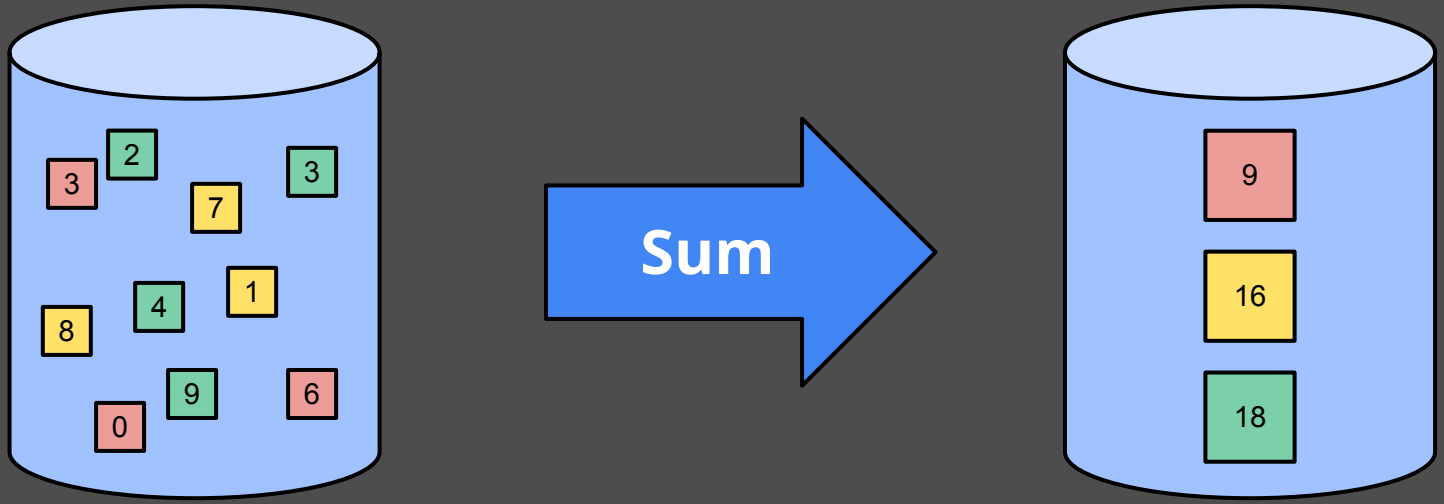
Input



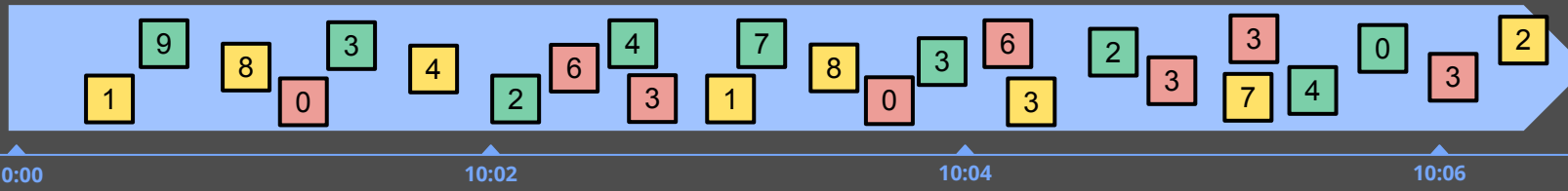
Output



Aggregation



Unbounded



Processing Time

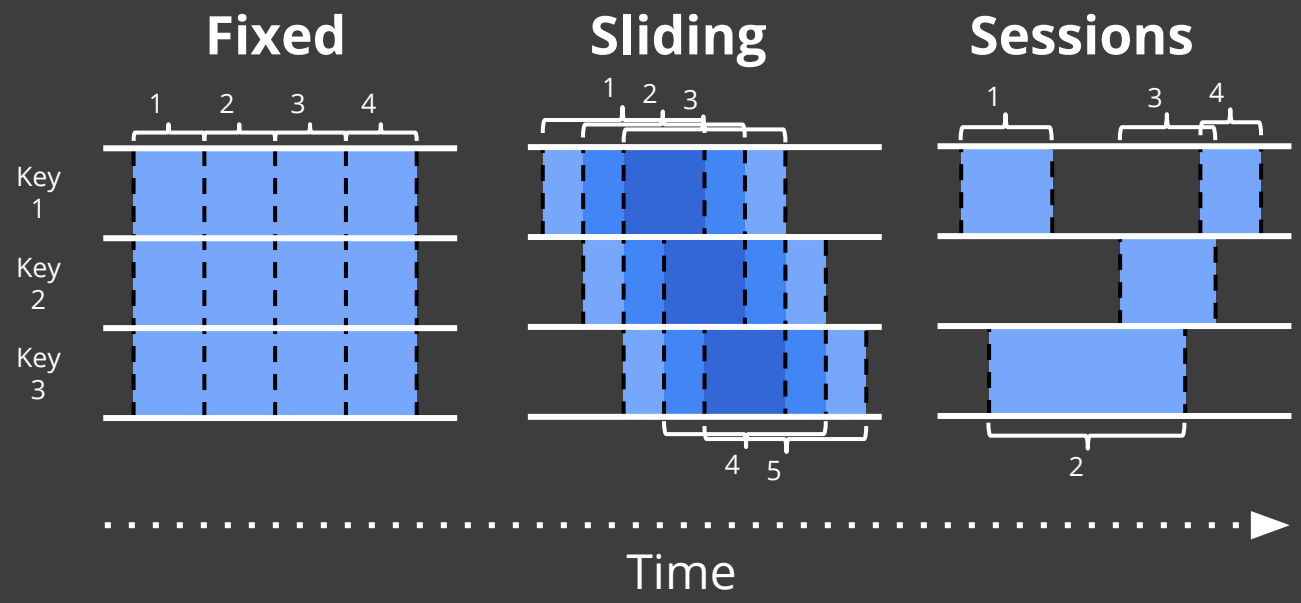
10:00 10:02 10:04 10:06

Sum

?

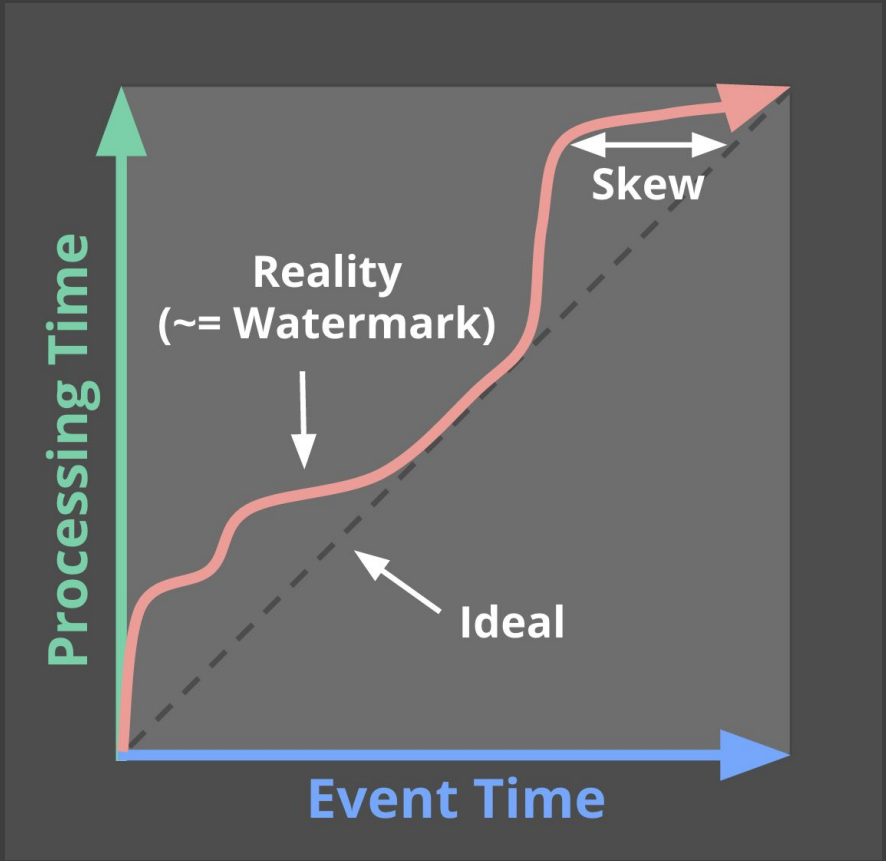
Where in event time?

Windowing divides data into event-time-based finite chunks.



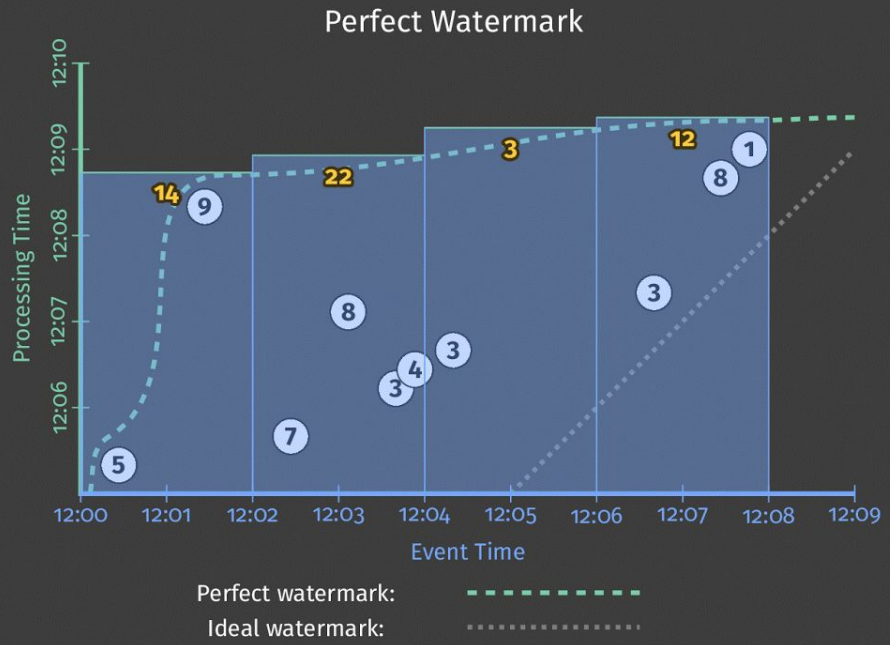
Often required when doing aggregations over unbounded data.

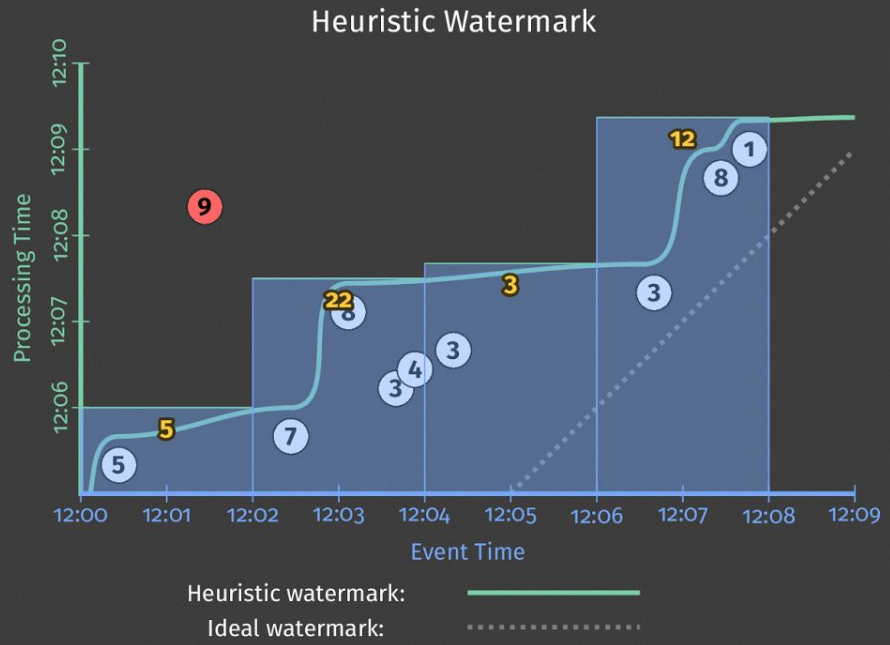
When in processing time?



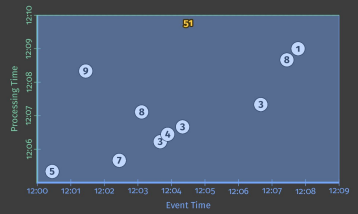
- Triggers control when results are emitted.
- Triggers are often relative to the watermark.

When: Triggering at the Watermark

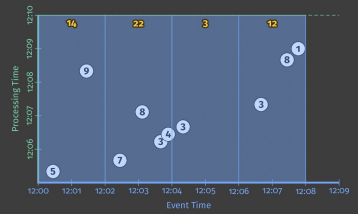




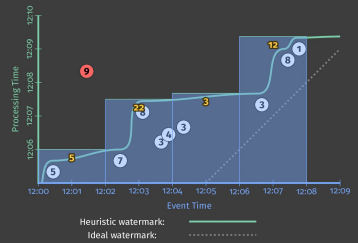
Customizing What When Where How



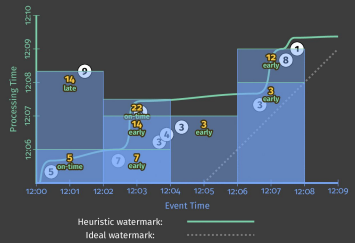
1.Classic Batch



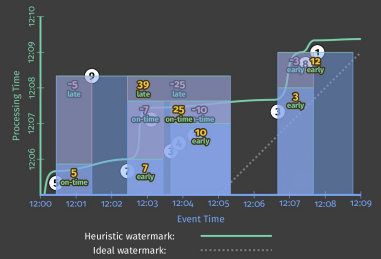
2. Batch with Fixed Windows



3. Streaming



4. Streaming with Speculative + Late Data



5. Streaming With Accumulations

The *Beam* Model & Cloud Dataflow

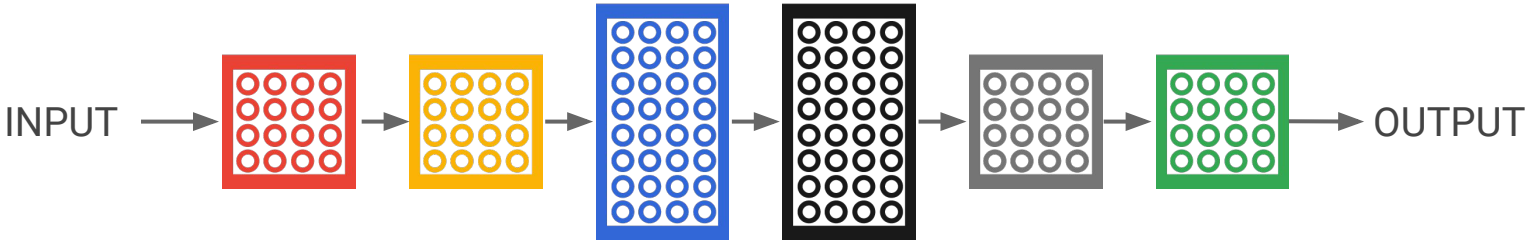
Apache Beam

a unified model for
batch and stream processing
supporting multiple runtimes

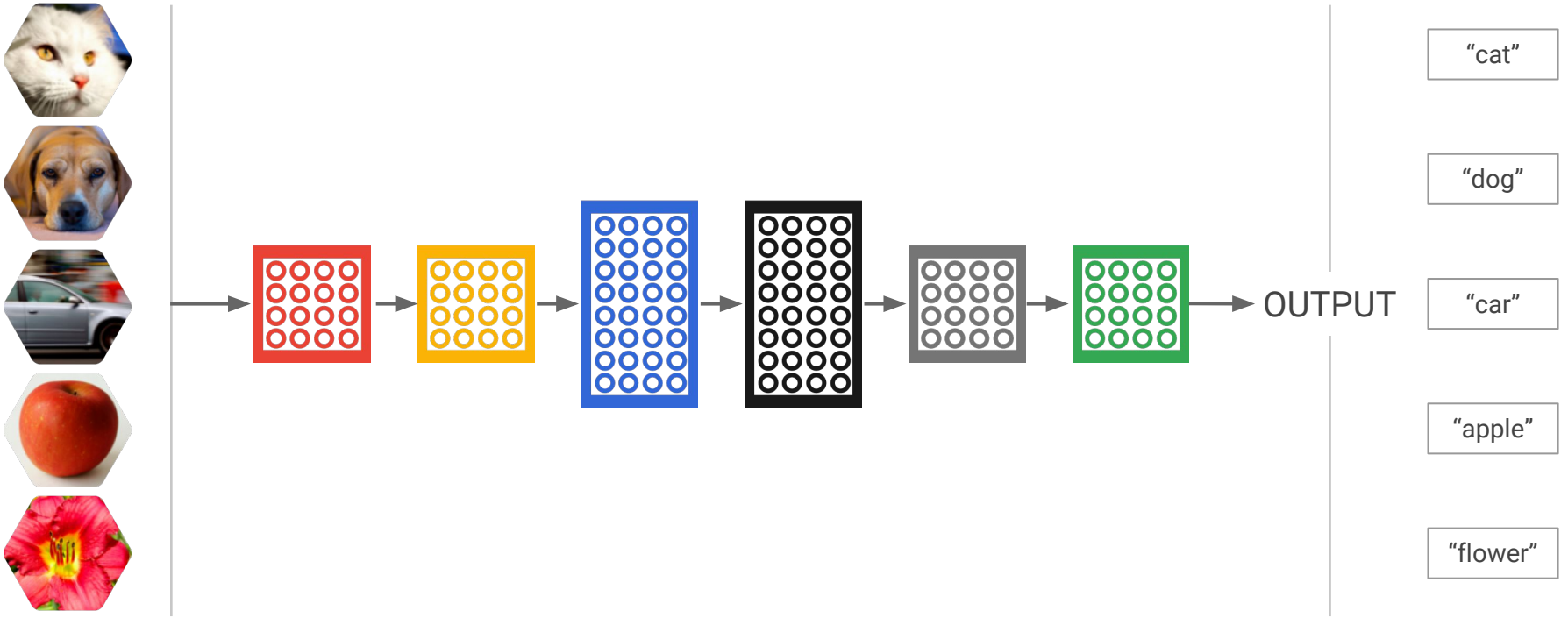
Google Cloud Dataflow

a great place to run Beam

one important technology we use is neural networks



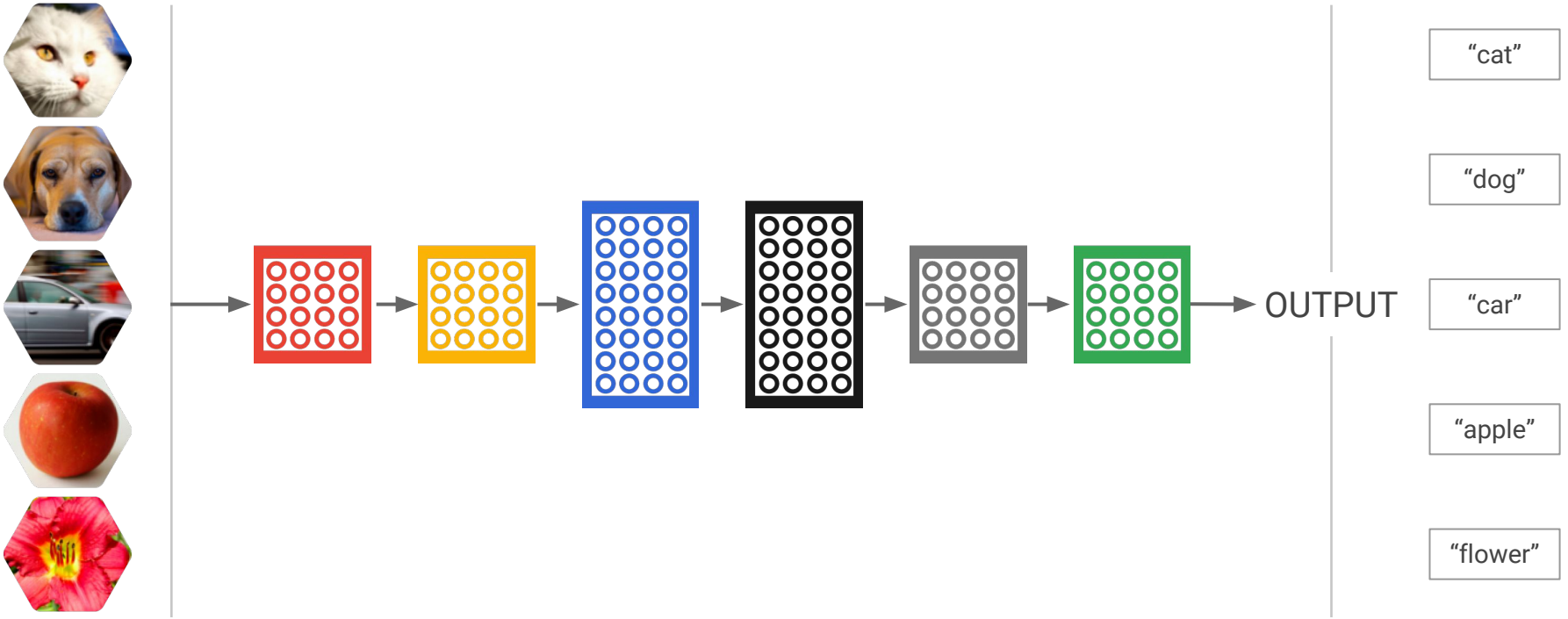
neural net models learn from examples



labeled photos

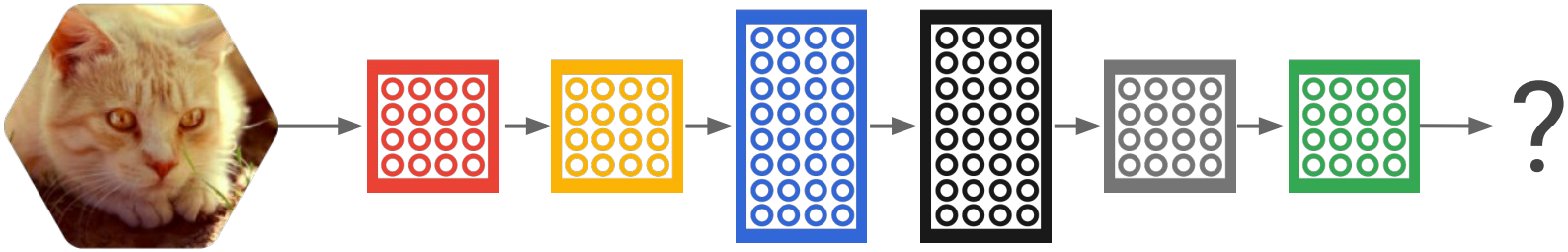
neural net models learn from examples

Make tiny adjustments to model so *output* is closer to *label* for a given image



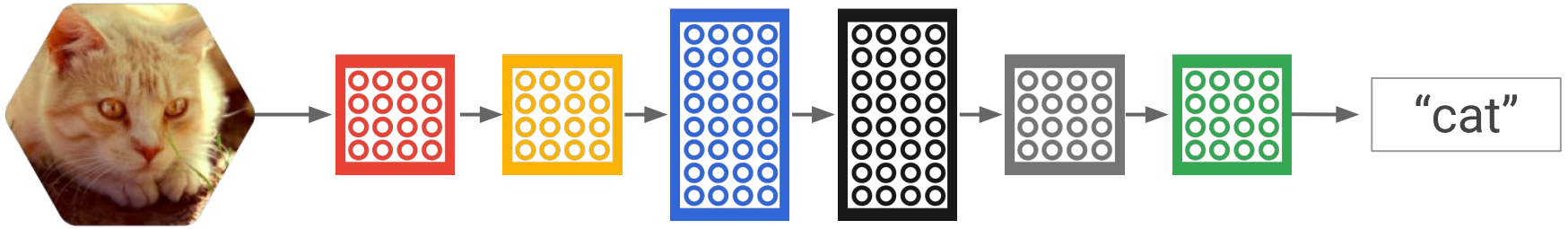
labeled photos

after a model is trained, you can test it



unlabeled photo

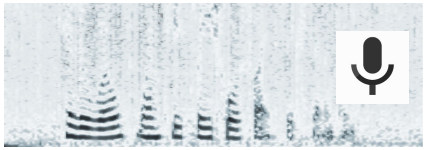
after a model is trained, you can test it



unlabeled photo

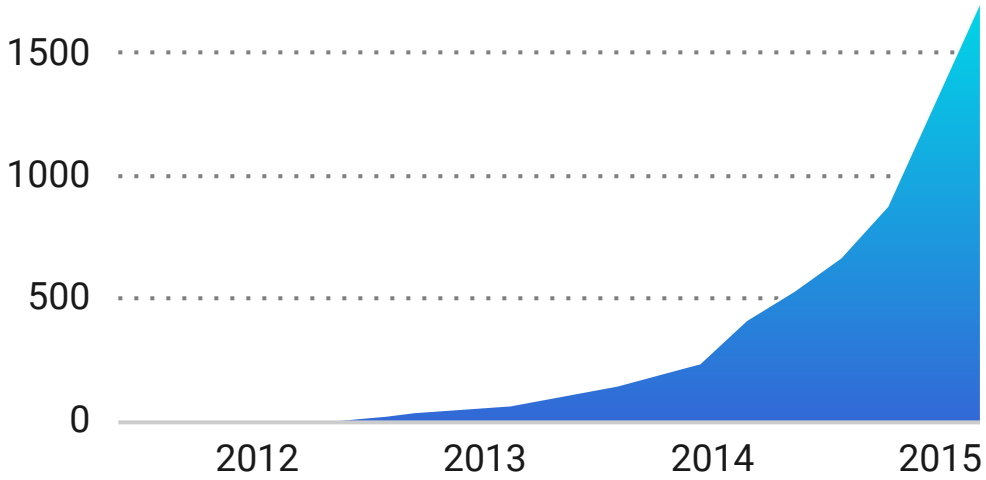
Can I Hug That?

powerful functions that neural nets can learn

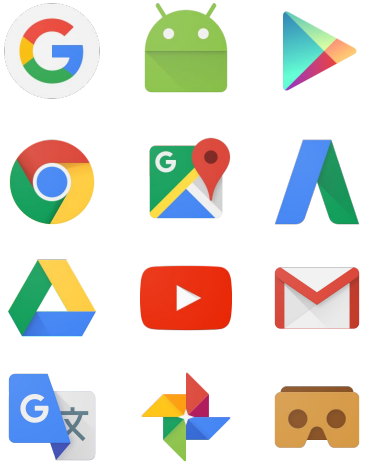
Input	Output
	“rice”
	“restaurants in Seoul”
안녕하세요	“hello!”
	“A close up of a small child holding a stuffed animal.”

Rapidly accelerating use of deep learning at Google

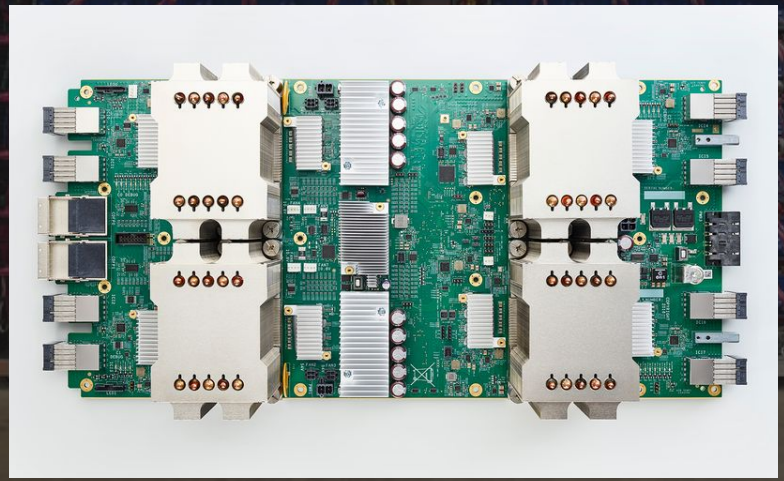
Number of directories containing model description files

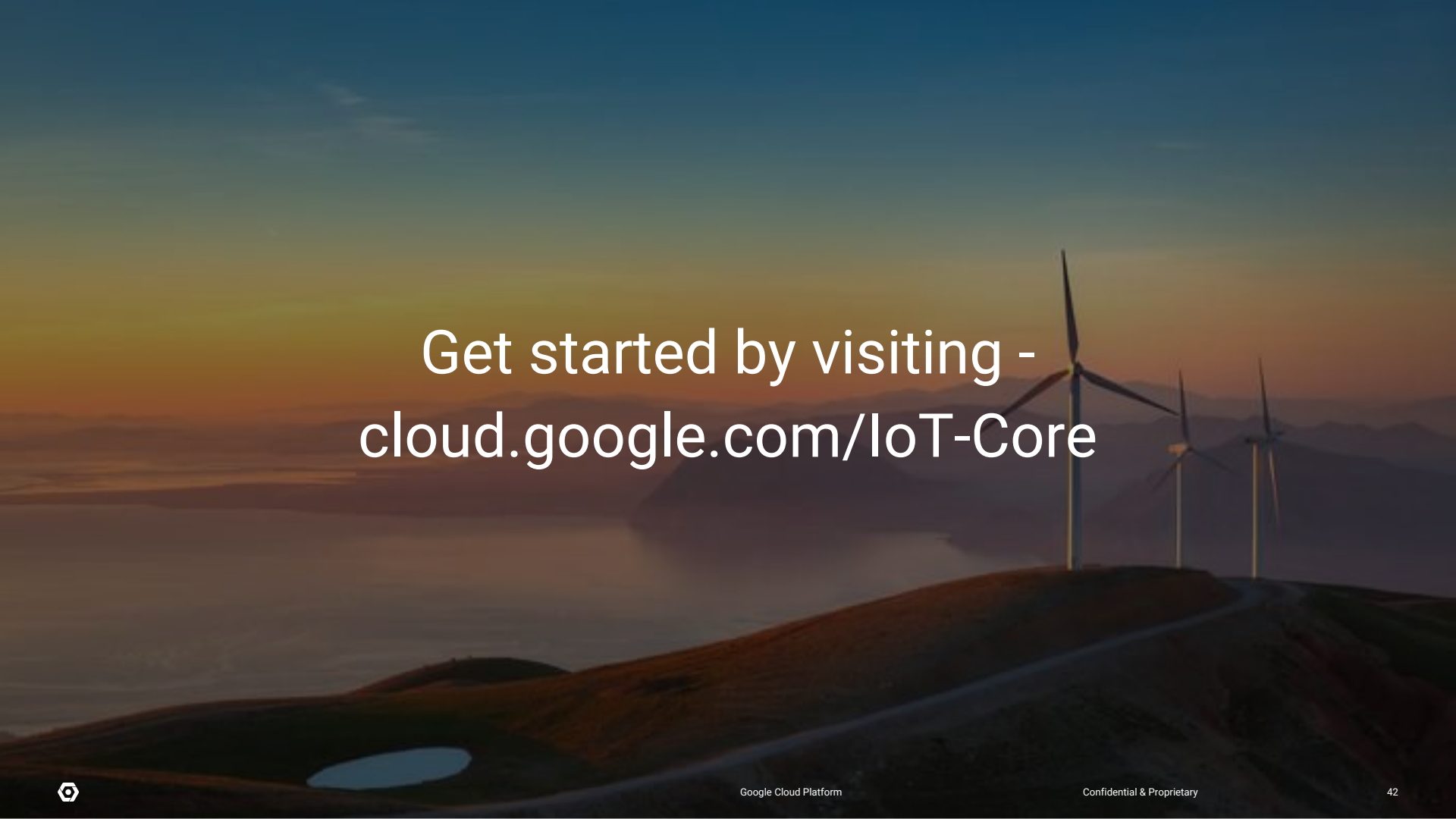


Used across products:



Our new Cloud TPU delivers up to 180 teraflops to train and run machine learning models.




Get started by visiting -
cloud.google.com/IoT-Core